Green Buildings

Jason Antal

2024-08-18

```
cat('To start our analysis off, we will look at buildings with lower than 10% occupancy
and compare them with the dataset to see if the previous analyst was justified in removing them.')
```

```
## To start our analysis off, we will look at buildings with lower than 10% occupancy
## and compare them with the dataset to see if the previous analyst was justified in removing them.
```

```
low_occupancy = filter(greenbuildings, leasing_rate < 10)

#Compare means for several predictors to see if any pattern stands out.
comparison <- data.frame(
  Low_Occupancy = round(sapply(greenbuildings[greenbuildings$leasing_rate < 10,
                         c("age", "Rent", "stories", "cluster_rent", "Gas_Costs", "size",
                           "class_b", "amenities", "green_rating")], mean),2),
  All_Buildings = round(sapply(greenbuildings[c("age", "Rent", "stories", "cluster_rent", "Gas_Costs",
                                          "class_b", "amenities", "green_rating")], mean),2)
)

print(comparison)
```

```
##              Low_Occupancy All_Buildings
## age                  54.42         47.24
## Rent                 22.44         28.42
## stories               4.82         13.58
## cluster_rent         23.99         27.50
## Gas_Costs             0.01          0.01
## size              62209.12     234637.74
## class_b               0.49          0.46
## amenities             0.12          0.53
## green_rating          0.00          0.09
```

```
cat('The table shows low-occupancy buildings are on average older, lower rent, less
    stories, are smaller, and are far less likely to have amenities. However, none of
    them are extreme outliers to the point we would want to omit them from our
    analysis. Instead of having "something weird going on", it is more likely
    that these are simply less desirable buildings that have lower occupancy
    rates as a result. We will include all buildings in this analysis.')
```

```
## The table shows low-occupancy buildings are on average older, lower rent, less
##     stories, are smaller, and are far less likely to have amenities. However, none of
##     them are extreme outliers to the point we would want to omit them from our
```

```
##     analysis. Instead of having "something weird going on", it is more likely
##     that these are simply less desirable buildings that have lower occupancy
##     rates as a result. We will include all buildings in this analysis.
```

```
cat("Let's now calculate how much more revenue a green building generates.")
```

```
## Let's now calculate how much more revenue a green building generates.
```

```
greenbuildings <- greenbuildings %>%
  mutate(cluster_factor = as.factor(cluster))

#Fit linear model
model <- lm(Rent ~ . - Rent, data = greenbuildings)

# Extract coefficient for green_rating
additional_rent_per_sqft <- coef(model)["green_rating"]

# Calculate total additional rent for a 250,000 sqft building
total_additional_rent <- 250000 * additional_rent_per_sqft

cat("Additional rent for green buildings: $", round(additional_rent_per_sqft, 2), "per square foot per y
```

```
## Additional rent for green buildings: $ 2.05 per square foot per year
```

```
cat("Total additional rent for a 250,000 sqft green building at 100% occupancy: $", round(total_additio
```

```
## Total additional rent for a 250,000 sqft green building at 100% occupancy: $ 511862.6 per year
```

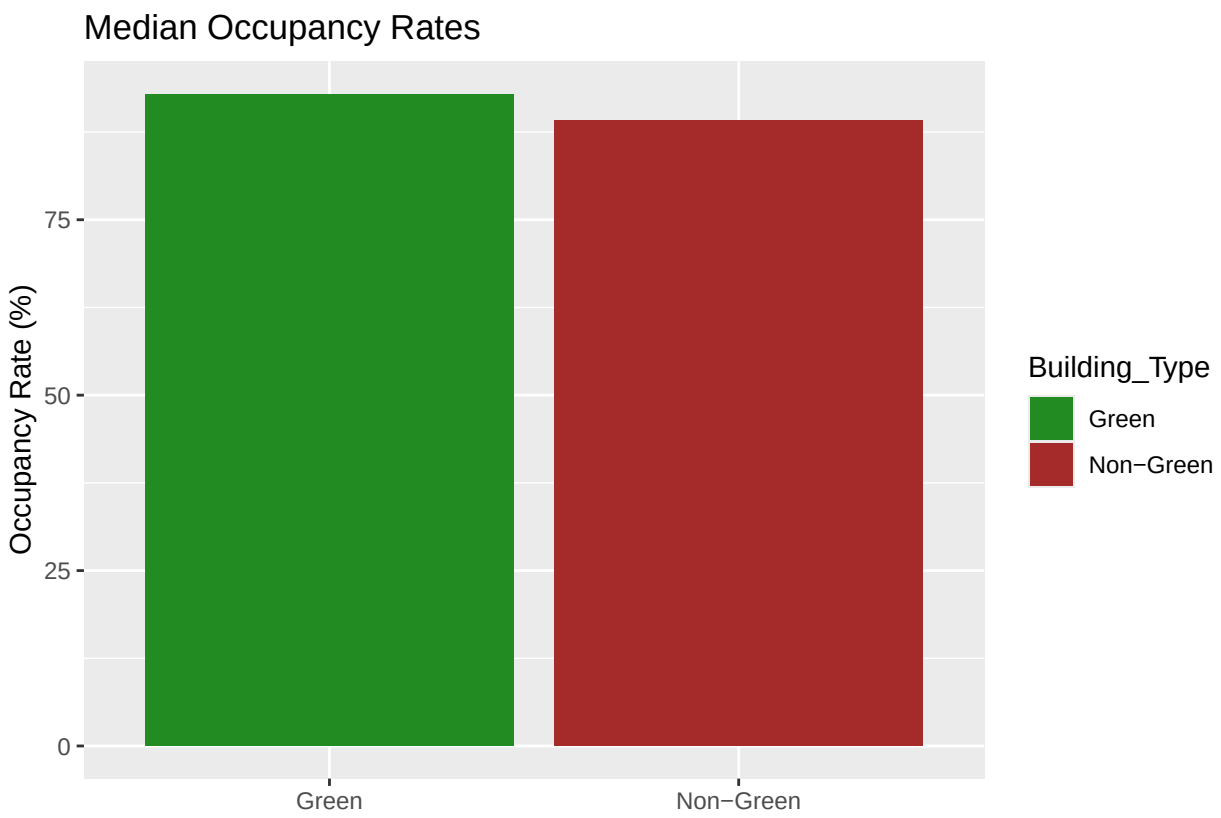
```
cat("This assumes 100% occupancy for green buildings, but it would be overly
    optimistic to rely on this assumption. Let's look at the median occupancy rate
    for green buildings instead and use that to determine exactly how much more
    money green buildings bring in when the expected occupancy rate is more realistic.")
```

```
## This assumes 100% occupancy for green buildings, but it would be overly
##     optimistic to rely on this assumption. Let's look at the median occupancy rate
##     for green buildings instead and use that to determine exactly how much more
##     money green buildings bring in when the expected occupancy rate is more realistic.
```

```
# Calculate median occupancy rates for green and non-green buildings
green_buildings <- filter(greenbuildings, green_rating == 1)
brown_buildings <- filter(greenbuildings, green_rating == 0)

med_green_occupancy <- median(green_buildings$leasing_rate)
med_brown_occupancy <- median(brown_buildings$leasing_rate)

ggplot(data.frame(Building_Type = c("Non-Green", "Green"),
                  Occupancy = c(med_brown_occupancy, med_green_occupancy)),
       aes(x = Building_Type, y = Occupancy, fill = Building_Type)) +
  geom_bar(stat = "identity") +
  labs(title = "Median Occupancy Rates",
       y = "Occupancy Rate (%)",
       x = "") +
  scale_fill_manual(values = c("Non-Green" = "brown", "Green" = "forestgreen"))
```

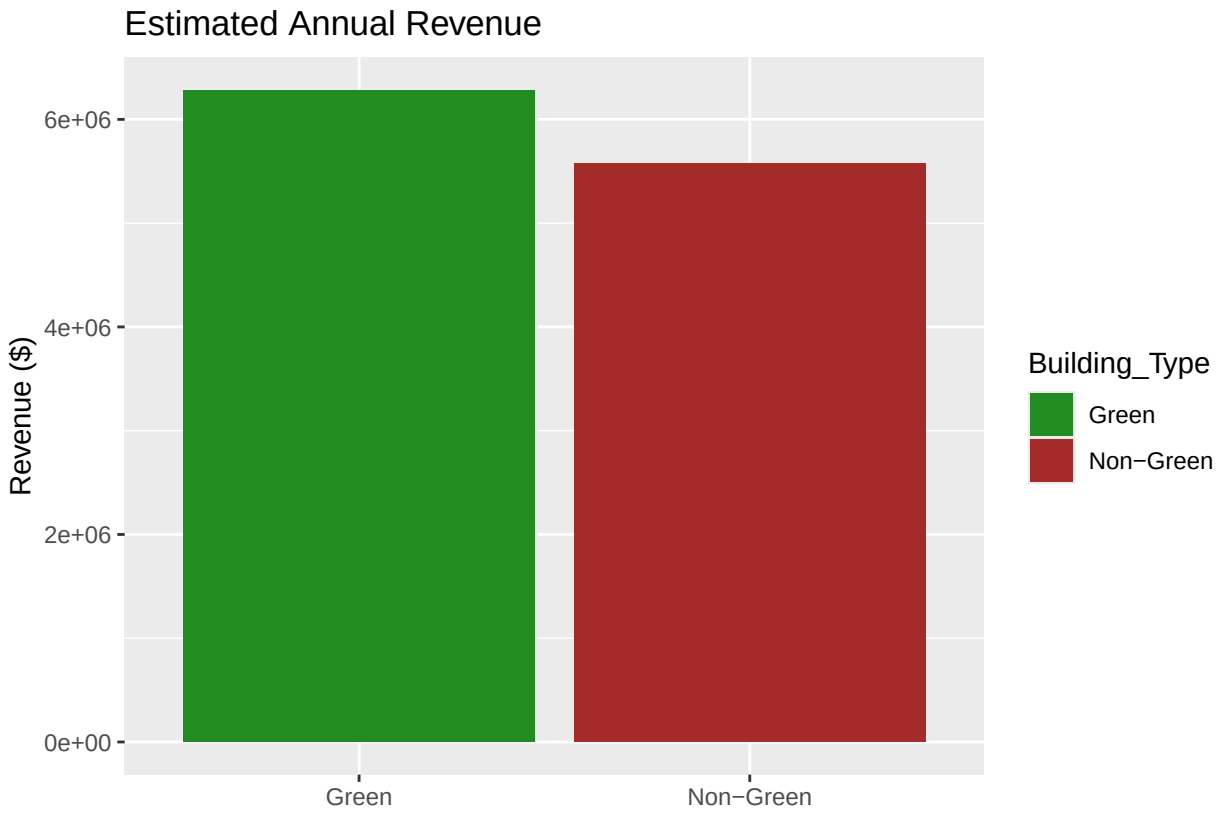



```r
# Calculate revenue for green and brown buildings
green_rent_sqft <- 250000 * (25 + additional_rent_per_sqft)
brown_rent_sqft <- 250000 * 25

avg_revenue_green <- green_rent_sqft * med_green_occupancy * 0.01
avg_revenue_brown <- brown_rent_sqft * med_brown_occupancy * 0.01

revenue_diff <- avg_revenue_green - avg_revenue_brown

ggplot(data.frame(Building_Type = c("Non-Green", "Green"),
                  Revenue = c(avg_revenue_brown, avg_revenue_green)),
       aes(x = Building_Type, y = Revenue, fill = Building_Type)) +
  geom_bar(stat = "identity") +
  labs(title = "Estimated Annual Revenue",
       y = "Revenue ($)",
       x = "") +
  scale_fill_manual(values = c("Non-Green" = "brown", "Green" = "forestgreen"))
```




```
cat("Additional annual revenue for green buildings: $", round(revenue_diff, 2), "\n")
```

```
## Additional annual revenue for green buildings: $ 709997.7
```

```
green_cost <- 5000000
payback_period <- green_cost / revenue_diff

cat("Estimated payback period:", round(payback_period, 2), "years\n")
```

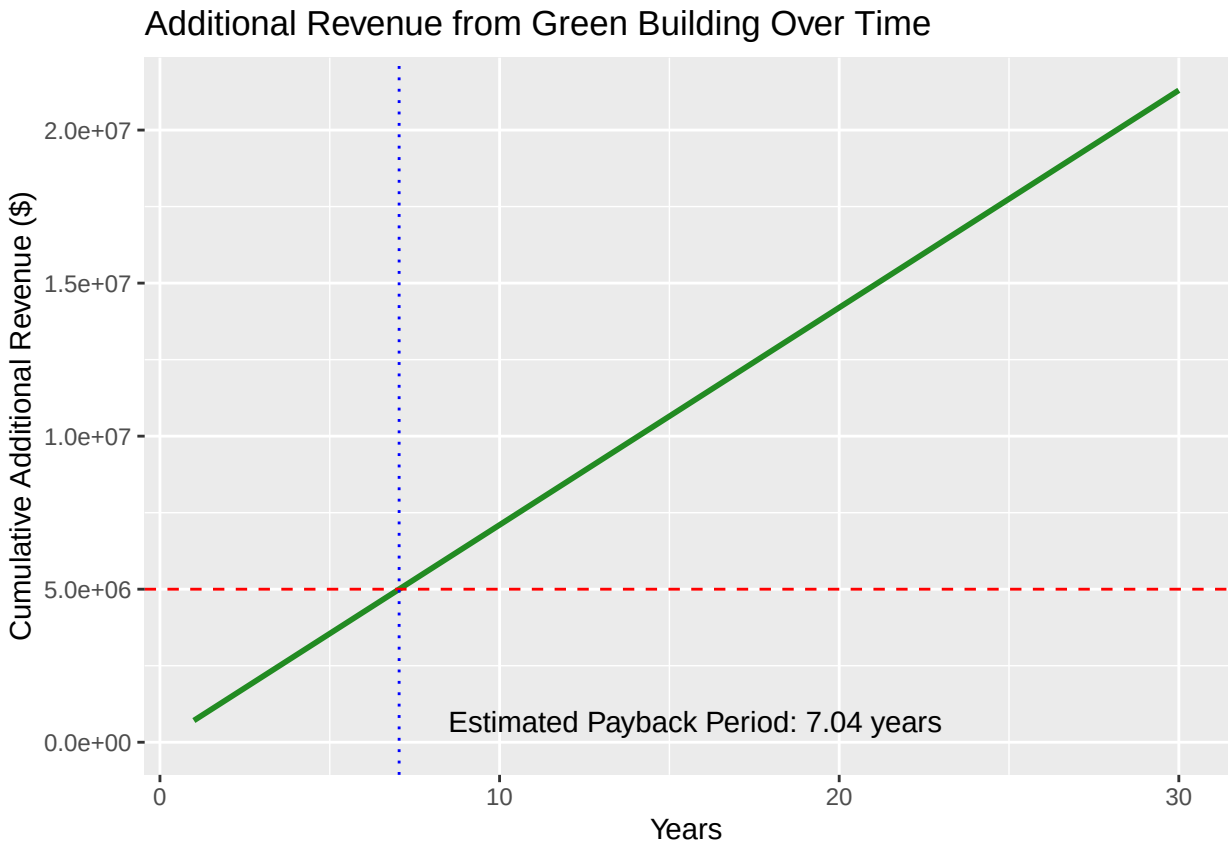
```
## Estimated payback period: 7.04 years
```

```
years <- 1:30
additional_revenue <- revenue_diff * years
investment_cost <- 5000000

ggplot(data.frame(Year = years, Revenue = additional_revenue)) +
  geom_line(aes(x = Year, y = Revenue), color = "forestgreen", size = 1) +
  geom_hline(yintercept = investment_cost, linetype = "dashed", color = "red") +
  geom_vline(xintercept = payback_period, linetype = "dotted", color = "blue") +
  annotate("text", x = payback_period, y = 0, label = paste("Estimated Payback Period:", round(payback_
           vjust = -0.5, hjust = -0.1) +
  labs(title = "Additional Revenue from Green Building Over Time",
       x = "Years",
       y = "Cumulative Additional Revenue ($)")
```

```
## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
```

```
## i Please use ‘linewidth‘ instead.
## This warning is displayed once every 8 hours.
## Call ‘lifecycle::last_lifecycle_warnings()‘ to see where this warning was
## generated.
```




```
cat("Even though your analyst's methods were not as accurate as a complete analysis,
    he was not too far off the mark. It would take only 7 years and 1 month to see
    a return on your investment if you chose to build a green building. In addition,
    there are other intangible positives not covered in this analysis, such as good PR.
    If you expect to be collecting rent on this building for longer than 7 years, and
    it makes sense for you financially, I would recommend building a green building.")
```